

By CATHERINE C. BLODGETT

# Duplication in Original Cataloging Among Six Research Libraries<sup>1</sup>

*Miss Blodgett is head of the Catalog Department, Providence Public Library.*

EXTENSION of cooperation in cataloging among libraries has been one of the most frequently advocated solutions for the "crisis in cataloging." In order to extend cooperative cataloging, several plans have been proposed. Among them are the development of a national union catalog suggested by Downs,<sup>2</sup> which would involve an elaborate system of regional union catalogs; and the somewhat complicated scheme advanced by Ellsworth,<sup>3</sup> whereby holdings of 1000 participating libraries would not only be recorded at the Library of Congress, but would also be circularized by means of book catalogs among the libraries themselves. The possibility of extending the services of the Library of Congress in cooperative cataloging has also been noted.<sup>4</sup> The utilization of the existing Union Catalog of the Library of Congress is involved in these proposals. The Library of Congress Cooperative Acquisitions Program had as one of its major features cooperative cataloging, based on special assignments, by the participating libraries.

Although these suggestions are on a na-

tional scale, it has been observed that there is also room for smaller, local schemes of cooperation, as for example the Connecticut Valley Project involving six college libraries of similar size and scope described by Rider.<sup>5</sup> As MacPherson<sup>6</sup> said in concluding her remarks on cooperative cataloging in 1936: "Duplication is still going on in many parts of the country, and even within different institutions of the same city."

Now a test of MacPherson's statement has been made for certain specific libraries within the New York metropolitan area. Within New York City, or no more than 100 miles from it, are five university libraries with facilities for graduate study: Columbia, New York, Fordham, Yale and Princeton. With these the New York Public Library ranks as a center for study by research workers. No attempt has been made as yet toward cooperation in cataloging among these libraries.

In his discussion of the Connecticut Valley Project Rider<sup>7</sup> noted:

Cooperative cataloging is, of course, not primarily a problem of cataloging rules or printing techniques, but rather one of organization and administration. Any first attempt at it must be tentative and exploratory.

As a first step in the direction of cooperation this study is offered. Its purpose is to determine if and in what ways co-

<sup>1</sup> Based on a master's essay, Columbia University, School of Library Service, 1948.

<sup>2</sup> Downs, R. B., ed. *Union Catalogs in the United States*. Chicago, American Library Association, 1942. p. 119-25, 283-315.

<sup>3</sup> Ellsworth, R. E. "Centralized Cataloging for Scholarly Libraries." *Library Quarterly*, 15:237-43, July 1945. See also his "Report." *Library of Congress Information Bulletin*, Nov. 16-22, 1948, Appendix.

<sup>4</sup> Haykin, D. J. "Way to the Future: Cooperative and Centralized Cataloging." *College and Research Libraries*, 3:156-62, 175, March 1942.

<sup>5</sup> Rider, Fremont. "Real Cooperative Cataloging." *Library Quarterly*, 13:99-112, April 1943.

<sup>6</sup> MacPherson, H. D. *Some Practical Problems in Cataloging*. Chicago, American Library Association, 1936. p. 124.

<sup>7</sup> *Op.cit.*, p. 99.

operative cataloging might be possible among six research libraries.

Recently the Library of Congress has been unable to supply printed cards for more than 50 per cent of the titles cataloged by the Columbia University Libraries. In response to a letter addressed to the Yale, Princeton, Fordham and New York University libraries and to the New York Public Library Reference Department, it was learned that cataloging statistics for 1946-47 from these libraries show a substantial percentage of titles for which no Library of Congress printed cards were available. It is assumed that duplication in cataloging of titles for which Library of Congress printed cards are not available is occurring among research libraries because of lack of co-operative procedures.

Fordham University estimated that printed cards were unobtainable for 25 per cent of titles cataloged. New York University reported that it was able to obtain cards for 65 per cent of its cataloging output; or, in other words, New York University was doing original cataloging for 35 per cent of its titles. At Yale University Library cataloging statistics show that in 1945-46 approximately 47 per cent of its books were cataloged without L.C. cards, and in 1946-47 about 40 per cent were so cataloged. At Yale statistics for serial cataloging would show a much higher percentage cataloged without printed cards. At Princeton the average for the two years, 1946-47, was 55.7 per cent of books given original cataloging. For the same period at the New York Public Library Reference Department 48.7 per cent of its titles, exclusive of serials and serial documents, were cataloged without L.C. cards.

To determine how much duplication in original cataloging was occurring among the six libraries—Columbia, New York University, New York Public Library,

Fordham, Yale and Princeton—a selected group of 500 main entries which received original cataloging at Columbia was checked in the catalog of each of the other five. These entries were chosen from material processed by the main Cataloging Department at Columbia in the spring and summer of 1946. Actual process slips were used for checking.

The sampling was made at random, but works in subject fields not thought to be common to all the libraries were excluded. Thus titles from the Schools of Law and Medicine at Columbia, and those titles specifically in the subjects of engineering and library science were eliminated. Entries for items in special collections were not included. Further, only entries in the Roman alphabet were included, which probably eliminated a number of titles in the Russian language. Proportions of various subject classes and languages were kept, as nearly as possible, representative of the kind and amount of work passing through the Columbia Cataloging Department.

In order to provide a basis for comparing the entries in the catalogs of the libraries, they were carefully examined point by point and were designated as included or lacking by checking the initials of the library. Such variations as difference in date of imprint, usually accepted at Columbia when present on L.C. cards, were allowed. In some cases duplication had to be assumed, as in comparison of the Columbia entries with those at Princeton, where preliminary pagination is not always given. In spite of careful checking in the L.C. Depository Catalog, a few L.C. printed cards did appear for titles checked in the other catalogs. The percentage of these was very low, probably not working out to more than about 1 per cent of the titles checked. They were not counted as duplicates.

Although the basic list was made up of

a random selection of 500 titles, they fell into four types: books, monographs, theses and publications of governments and other corporate bodies. Serials were excluded from consideration because they form so large and heterogeneous a class of material, and because libraries tend to differ in their treatment of them.

**BOOKS:** Of the four categories in the list the largest was that of books, represented by 400 of the 500 titles. Books were taken to include all works which did not fall naturally into any of the other classes. The majority of works cataloged at Columbia are in this class.

**MONOGRAPHS:** The differentiation among books, theses and publications of corporate bodies is fairly obvious, but that between books and monographs was arbitrarily defined for this study. For the purpose of this study any nonliterary work of less than 100 pages on a single subject, which did not come naturally within the scope of theses or publications of corporate bodies, was considered a monograph. Monographs were the second type of material in size. Of the total 500 titles there were 73 of them.

**THESES:** Since kinds of material not common to all libraries under investigation were to be kept at a minimum, fewer theses were selected than might have been. It was assumed that the New York Public Library would have fewer of these than the university libraries would. Only 20 titles on the list were theses, although Columbia catalogs a far greater number of them than is here represented.

**GOVERNMENT AND CORPORATE PUBLICATIONS:** The elimination of serials from the study automatically eliminated many government documents and publications of corporate bodies. There were only seven of the titles in this class, or a little over 1 per cent of the total.

**SUBJECTS:** The Dewey classification as used at Columbia was chosen for the subject division of the list. Literature ranked first among subjects, with a total of 162 entries. It was followed by history with 97; social sciences, 93; fine arts, 49; science, 32; philosophy, 30; religion, 17; technology, 12; general class, five; comparative philology, three.

**LANGUAGES:** English led all the other languages on the list, with 258 titles. German was second, with 67. Spanish and Portuguese considered together had 63, and French 49. There were 21 entries in the three Scandinavian languages, and ten, nine, three and three in Dutch, Italian, Russian and Latin, respectively. A scattering of other languages appeared on the list. These figures approximate the proportions of languages in the Columbia Cataloging Department's output. An exception should be made of Russian, which at Columbia accounts for a larger proportion of titles than is here indicated. The cause of this discrepancy, as stated above, is that this study eliminated titles in the Russian alphabet.

**DATES:** In order to determine the extent of the currency of the material being cataloged by the various libraries, three time periods of unequal length were set up. These were: 1800-1899, 1900-1939 and 1940-1946. These periods were chosen because the majority of books cataloged at Columbia were published in the twentieth century, and of all titles considered, the greatest number had appeared since 1940. That this time scheme was essentially suitable is borne out by the fact that 294 of the entries fell within the 1940-1946 period. One hundred and thirty-six were published between 1900 and 1939. Only 70 appeared before 1900.

These titles, with the exceptions noted, were taken as typical of Columbia's original cataloging production during 1946. Since



they were typical of a large library's cataloging, it was held that this might properly be used as the instrument to test the assumption that there is duplication in cataloging among a group of libraries in a certain locality.

The data collected in the investigations at the New York Public Library and at the libraries of Yale, Princeton, Fordham and New York University were studied according to the method set up for the analysis of the basic list, and the results obtained at each library were compared with that list.

Most duplication with the list from Columbia occurred in the catalog of the New York Public Library. This library contained 94 of the entries, or slightly less than 20 per cent. At Yale 86 of the entries were found, representing 17.2 per cent of the Columbia list. Approximately 13 per cent, 65 of the 500 titles, were found in the Princeton University Library. New York University had only 13 of the titles checked, and Fordham had 12, or fewer than 3 per cent in each case.

Distribution of the duplicate entries by type of material, subject, language and time period followed closely the distribution of the entries on the basic list in the check of all five libraries. It was generally found that most duplications occurred in the fields of history, literature and the social sciences; that books were the form of material most frequently occurring; that English was the

language of most titles; and that most were published in the period 1940-1946. Of the 20 theses on the basic list 13 were found at Yale and 10 at Princeton.

The percentage of duplication revealed by the study does not suggest a general program of cooperative cataloging among the six libraries. The possible exception to this conclusion is cooperation between Columbia and the New York Public Library. A further study of the libraries shown here to have the most in common (Columbia, New York Public Library, Yale and Princeton) might indicate some possibility of cooperation along lines more restricted by subject, language, type of material or date of publication than was the list which was the basis of the study. For example, there might be cooperation on the basis of specialization in acquisitions programs. This idea has been given impetus by the projects developed during the war, such as the library of Congress Cooperative Acquisitions Program.

How many of the items included in the list of 500 titles are held by the Library of Congress and not cataloged was an unknown quantity. It is apparent, however, that so long as the Library of Congress cannot produce its cards quickly, American libraries will need to prepare some of their own cataloging. While this situation obtains, the possibilities of cooperative cataloging among groups of libraries should be explored.

## Kentucky Archives Fellowship

An Archives Fellowship will be offered by the Margaret I. King Library, University of Kentucky, for the fiscal year 1951-52 or for the academic year. Stipend is \$100 per month with part-time work assignments under direction of University of Kentucky Archivist. Projects will be related to fellow's interests if possible. Partial loads may be carried in any academic department, and fellow may become a candidate for a degree. Fellowship is renewable. Apply to director of libraries, University of Kentucky, Lexington, Ky.